# Assessment of the impact of social network and seed exchange on crop genetic diversity

D4-3-1

*Mathieu THOMAS, Abdel Kader NAINO JIKA, Isabelle GOLDRINGER*

*2019-08-29*

## Contents

# List of Figures

# 1 Introduction

## 1.1 General context

Sustainable farming in the context of global change should be based on diversified agro-ecosystems, including among and within variety diversity, since they usually provide more resilience in face of fluctuating environmental conditions. Seed circulation among actors (farmers, gardeners, mainteners ...) is one of the key elements that influence diversity dynamics over time. **The project's aim is to use a simulation based approach to numerically analyse the impact of seed circulation on crop biodiversity dynamics and explore different scenarios of crop diversity management.** For this, we developed CropMetapop, a software allowing for simulating the genetic evolution of a crop metapopulation including the different evolutionary forces (genetic drift, selection, mutation and migration) and demographic processes (extinction, colonization and population growth). The main difference with already existing softwares that have been designed for natural metapopulations is that circulation of seed lots among actors (seed exchange) has been introduced. For that, a dedicated migration and/or colonization of seed samples among populations was defined. The model has been designed and the code implemented in Python by using the SimuPop library.

## 1.2 Objective

The objective is to use the CropMetaPop model to explore the impact of different modes of organization of Community Seed Systems (CSS), on genetic diversity and population differentiation and adaptation. This may allow to identify parameters that are most influential on crop diversity dynamics and explore scenarios of evolution (breeding practices, seed circulation,...) for CSS. Relevant seed circulation networks and scenarios of evolution will be co-designed with interested DIVERSIFOOD partners in a participatory approach and numerically assessed with the simulation model. The aim is that the discussion and co-design of scenarios among the modelling team and the CSS partners help them reflect on their pratices by bringing another point of view. Moreover, the objective is to produce opinions and recommendations resulting from approaches based on modelling in population genetics and shared between scientists and CSS actors.

## 1.3 Approach

We use an iterative process where the functioning of the model and results for simple cases are first shown to CSS partners for discussion. This should lead to new scenarios proposed, that account for a more realistic functioning of the CSS. Then, these new scenarios are implemented with modified values of the parameters and the results of simuations are further discussed among the modelling team and the CSS partners. During DIVERSIFOOD annual meeting in February 2017, CropMetaPop functioning was shown using three sets of parameters for mating system and populations size meant to represent different crop species. These three types of crops were tested with three modes of seed circulation : no seed exchange at all, seed exchanges according to a random network and seed exchanges according to a centralized network. This was used as a first basis for discussion among the CSS partners and the modelling team. Based on the comments, the simple case studies were modified and completed to allow for a broader understanding of the impact of the parameters accounted for (mating system, population size and seed circulation network). In addition, four CSS partners (RSP tomato and wheat groups, RSR wheat and RAS tomato groups) volonteered to participate in a scenario co-design exercise. During DIVERSIFOOD annual meeting in March 2018, 12 case studies were presented and further discussed with all partners as well as the first propositions for modelling the four real CSS functionings. The results of the case studies presented at this workshop and the discussions make the basis for Section 3 of this deliverable. The co-design work conducted with the four CSSs has continued during 2018 and until the last annual meeting in January 2019; this is presented in the Section 4 of the deliverable.

### 1.4 Content of the deliverable

In the following we first describe the approach, how the simulation model works and what outputs of the model we considered (i.e. the demographic and genetic parameters used to analyse the response) (Section 2). Then we present 20 contrasting case studies to explore the effect of (i) a biological factor of the crop species considered (mating systems = outcrossing or mainly selfing), (ii) a farmer management practice (population sizes = small or large) and (iii) a social organization aspect (network of seed circulation = none, decentralized, centralized or community) (Section 3). Finally, Section 4 shows the work conducted with the four CSSs (RSP tomato and wheat groups, RSR wheat and RAS tomato groups), from understanding their functioning to co-designing scenarios of crop diversity management and assessing the impact of some key factors (environmental conditions and farmers' practices, seed circulation network, extinction and colonization modalities,...).

## 2 Modelling community seed systems

### 2.1 Model hypotheses

CropMetaPop is an individual-based model which allows for simulating the genetic evolution of several populations over time. This model belongs to the theoretical framework of the metapopulation but it also takes into account crop management by farmers' organizations. A **crop metapopulation** in this context may for instance correspond to different versions, named **crop populations**, of the same population-variety. One generation corresponds to a life cycle from seed sowing to plant harvesting.

#### 2.1.1 Demographic hypotheses

Each crop population is composed of a finite number of **crop plants**.

These crop populations are grown in a finite number of **farms**. Each farm is characterized by the maximum number of plants that can be grown on it and named **carrying capacity**.

The number of plants of a crop population, named the **demographic size**, evolves according to the number of offspring produced per individual per generation, named **fecondity**, which is species dependent. The demographic size of a crop population can grow up to the carrying capaticity of the farm in which the population is grown.

Each crop population can randomly disapear through an **extinction** process. Thus, the corresponding farm becomes empty. Extinction can correspond to different situations such as climatic or pest disasters, or other difficulties encountered by a farmer to maintain his population to the next generation.

After an extinction event, the empty farm can be randomly filled, through a **colonization** process, corresponding to seed circulation from one or from several farms still growing the population-variety. Even though it is a random process, seed can circulate from one farm to the other only if farmers of these farms belong to the same social network. We thus make the hypothesis that **farmers' seed network are embedded in farmers' social network**. Farmers' social network is also named **community seed system**.

#### 2.1.2 Genetic hypotheses

Each crop metapopulation is characterized by a mating system ranging from open-pollination to self-pollination.

Each crop plant consists in a finite set of genetically independant loci.

Each locus is biallelic or multi-allelic. Mutation may occur with a certain rate at each reproduction event.

One locus corresponds to a neutral marker or a maker located in a gene region associated to a selective value. Similar or different optimums for plant fitness can be assigned to the farms in order to apply selection to

local populations. Each crop population is submitted to both selection and genetic drift depending on its demographic size.

## 2.2 Technical features

### 2.2.1 Software

CropMetaPop is based on simuPOP (http://simupop.sourceforge.net), a python library for the simulation of populations. CropMetaPop is a console program writen in python2.7 using an object oriented approach. CropMetaPop can be used on any computer platform. Limits of CropMetaPop are the same than simuPOP (http://simupop.sourceforge.net/Main/FAQ#toc4).

CropMetaPop is a free software under the terms of the version 3 of the GNU General Public License published by the Free Software Foundation.

### 2.2.2 Input and output data

Input: CropMetaPop is launched using a settings file. The settings file is a text file with flexible and friendly structure.

Output: CropMetaPop produces the raw results of the simulation. These results summarize information on genotypes for each replicate, generation and population, such as the number of individuals for each mono-locus genotype or multi-locus genotype or haplotype. Output also proposes history of every seed lot transfer events (colonization).

## 2.3 Output statistics

Input parameters need to be gathered before any simulation. Then, simulation produces genetic composition of each population at each generation for each replicate. It is thus necessary to compute demographic and genetic statistics to summarize the output in order to understand the response of the model for each scenario over time.

### 2.3.1 Demographic statistics

In the following sections, a crop population is by definition alive. When a crop population extincts, then the farm hosting the crop population becomes empty, without any population.

**Survival rate:** $S(t)$

Survival rate ($S(t)$) is calculated as the ratio of the number of crop populations at time $t$ over the total number of farms. $S(t)$ ranges from 0 (when all farms are empty) to 1 (when all farms are occupied). The formula to calculate $S(t)$ is detailled in Appendix 6.1.

**Occupancy rate:** $O(t)$

The occupancy rate $O(t)$ is calculated as the average of the occupancy rate of the crop populations composing the metapopulation at time $t$. The occupancy rate of a crop population at time $t$ ranges from 0 to 1 and corresponds to the ratio between the number of crop plants composing this population and the carrying capacity of this population. The formula to calculate $O(t)$ is detailled in Appendix 6.2.

#### 2.3.2 Genetic statistics

**Genetic diversity:** $H_s(t)$

The genetic diversity ($H_s(t)$) is calculated as the average of the expected heterozygosity of all crop populations composing the metapopulation at time $t$. In each population, the expected heterozygosity ranges from 0 (no diversity at all in the population, *i.e.* all plants in the population are genetically similar) to 1 (all plants in the population are genetically different), thereby leading to $H_s(t)$ also varying between 0 (no diversity in all populations) and 1 (maximum diversity in ll populations). The formula of the genetic diversity and of the expected heretorozygosity are detailled in Appendix 6.3.

**Genetic differentiation:** $F_{ST}$

The global level of divergence among populations composing the metapopulation is obtained with the differentiation index $F_{ST}$. The general formula to calculate $F_{ST}$ is detailled in Appendix 6.4. $F_{st}$ ranges from 0 to 1. It allows us to evaluate the level of differentiation among populations from different farms. A 0 value for $F_{ST}$ means that the different populations have a very similar genetic composition while a $F_{st}$ of 1 indicates that the different populations have very different genetic compositions.

# 3 Theoretical scenarios

## 3.1 Scenarios presentation

### 3.1.1 Scenario and simulation definitions

In this document, a scenario consists of a unique combination of CropMetapop input parameters. We move from one scenario to another by changing at least one parameter value. Note that some parameters are fixed for all scenarios: the number of generations, the number of replicates, the number of populations, the number of markers, the number of alleles, the mutation rate and the fecundity. Extinction and colonization rates are fixed for all scenarios involving extinction and colonization.

In this document, a simulation is the numerical assessment by CropMetaPop of one scenario, *i.e.* a particular set of input parameter values is used by CropMetapop to make a crop metapopulation evolve numerically over several generations.

CropMetapop simulations take into account the different evolutionary forces as well as farmers' management practices and the impact of their social organization through seed circulation. however, only the mating system and the population size of the populations are considered as farmers' management practices, while the structure of the seed circulation network is summarized by graphs.

### 3.1.2 Farmers' management practices modelling

#### 3.1.2.1 Definition of the management practices

The management pratices applied by farmers are represented in these theoretical scenario by the mating system and the carrying capacity. As we always simulate populations at there carying capacity, in the following of the document, population size is strickly equivalent to carrying capacity.

We consider four situations of management practices that can be divided into two categories (Fig. 1):

- Crop metapopulations cultivated in large fields (large population size with a carrying capacity of 10000 individuals), including an outcrossing species (such as maize, with a selfing rate of 0) and a self-pollinated species (using a high selfing rate of 0.95, like wheat).

- Crop metapopulations grown in small plots (small population size with a carrying capacity of 100 individuals), including an outcrossing species (such as cabbage, with a selfing rate of 0) and a self-pollinated species (using a high selfing rate of 0.95, like tomato).

Note that CropMetaPop only simulates evolving populations. The taxonomic level of species does not exist explicitly in the model. We introduce the concept of species here only for didactic reasons.

| Mating system | Population size | Example | |
|---|---|---|---|
| Outcrossing (s=0) | Large (n=10000) | Maize |  |
| Selfing (s=0.95) | Large (n=10000) | Wheat |  |
| Outcrossing (s=0) | Small (n=100) | Cabbage |  |
| Selfing (s=0.95) | Small (n=100) | Tomato |  |

Figure 1: Illustration of the four situations of management practices assessed by simulation. $s$ corresponds to the selfing rate and $n$ to the carrying capacity

#### 3.1.2.2 Initialization phase

The objective of the initialization phase is to generate populations close to genetic equilibrium with similar features for their genetic diversity and corresponding to each of the two mating systems (selfing and outcrossing). Starting all simulations from these initial populations allows to obtain results that will not depend from a particular initial composition of the populations. For that, we started with a large population (10000 individuals) closed to Hardy Weinberg equilibrium with a high level of diversity (allelic frequencies of 0.5 for each of the 10 bi-allelic loci) that was then conducted for 100 generations under either selfing (selfing rate of 0.95) or outcrossing (selfing rate of 0) with 10000 individuals each.

At generation 100, we sampled a large population of 10000 individuals and a small population of 100 individuals in both selfing and outcrossing populations. The genotypic frequencies for each population were calculated to check that a stable state was reached and was consistent with theoretical expectations. These samples were then used to initialize any simulation.

### 3.1.3 Scenarios without seed circulation

First, the four situations of management practices are assessed without any seed circulation. For each of the four cases, introducing a probability of extinction was compared to no extinction. We obtain 8 scenarios (scenarios #1 to #8) to compare. They are summarized in Table 1.

### 3.1.4 Scenarios with seed circulation

The four situations of management practices are also assessed with seed circulation. A network structure was used to represent social relationships among farmers/farms. A network is composed of a set of nodes that corresponds to the farms, and to a set of links (edges) that represent the social relationship between two farmers/farms that can allow a seed flow between two farms. Different network structures, also named topologies, were used to represent the diversity of social organizations (scenarios #9 to #20) (Fig. 2).

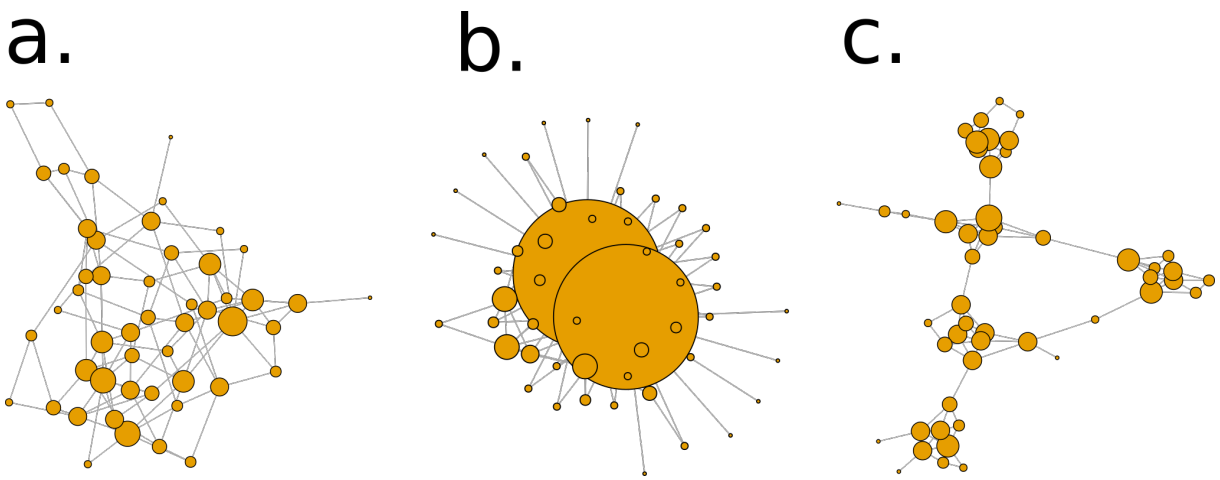


Figure 2: Three types of social organizations: a. decentralized network, b. centralized network, c. community network

Three types of social organizations have been investigated :

- a. A decentralized network that represents a situation where everyone has the same social status. This means that, on average, everyone has the same number of social relationships.
- b. A centralized network that represents a situation where a few actors are at the centre and through which all information circulates. On the other hand, most of the actors are peripheral with very few social links. This network could represent a case where a few community seed banks provide seeds to many farmers.
- c. A community network that represents a situation where actors are organized into sub-groups, like local organizations, and exchange more information within their group than between groups.

We obtained 12 additional scenarios by crossing management practices with social network structures (scenarios #9 to #20, parameter values summarized in Table 2).

Table 1: Parameter values used by CropMetaPopulation for the scenarios without seed circulation

| Parameter/scenario | #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 |
|---|---|---|---|---|---|---|---|---|
| Number of generations | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 |
| Number of replicates | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| Number of populations | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Carrying capacity | 100 | 100 | 100 | 100 | 10000 | 10000 | 10000 | 10000 |
| Number of marker | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| Number of allele | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| Mutation rate | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Fecundity | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| Selfing rate | 0.95 | 0.95 | 0 | 0 | 0.95 | 0.95 | 0 | 0 |
| Extinction rate | 0 | 0.1 | 0 | 0.1 | 0 | 0.1 | 0 | 0.1 |
| Colonization rate | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Table 2: Parameter values used by CropMetaPopulation for the scenarios with seed circulation

| Parameter/scenario | #9 | #10 | #11 | #12 | #13 | #14 | #15 | #16 | #17 | #18 | #19 | #20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of generations | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 |
| Number of replicates | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| Number of populations | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Carrying capacity | 100 | 100 | 10000 | 10000 | 100 | 100 | 10000 | 10000 | 100 | 100 | 10000 | 10000 |
| Number of marker | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| Number of allele | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| Mutation rate | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Fecundity | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| Selfing rate | 0.95 | 0 | 0.95 | 0 | 0.95 | 0 | 0.95 | 0 | 0.95 | 0 | 0.95 | 0 |
| Extinction rate | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| Colonization rate | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| Colonization network | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |

## 3.2 Results

### 3.2.1 Demographic results

#### 3.2.1.1 Results without seed ciculation

In the absence of seed circulation (colonization rate equal to zero) and without extinction (scenarios **#1, #3, #5** and **#7**), the **survival rate**, which is the proportion of non-empty farms, remained unchanged over generations whatever the selfing rate and the carrying capacity. When we introduced extinction processes but still without colonization (scenarios **#2, #4, #6** and **#8**), the survival rate logically decreased over time to tend towards 0, which corresponds to the extinction of the metapopulation (Fig. 3A, results are only shown for scenarios **#5** and **#6** as the results remained the same for both large and small populations size and selfing and outcrossing crop populations).

Unlike the **survival rate**, the **occupancy rate** remained the same over generations even when extinction is introduced (Fig. 3B, results are only shown for scenarios **#5** and **#6** as the results remained the same for both large and small populations size and selfing and outcrossing crop populations). Indeed, as initialization of the simulations was done at the carrying capacity, all farms that are not empty, are occupied by populations at their carrying capacity and the occupancy rate is estimated only based on these populations. As expected, for both large population sizes and small population sizes, the results remained the same.

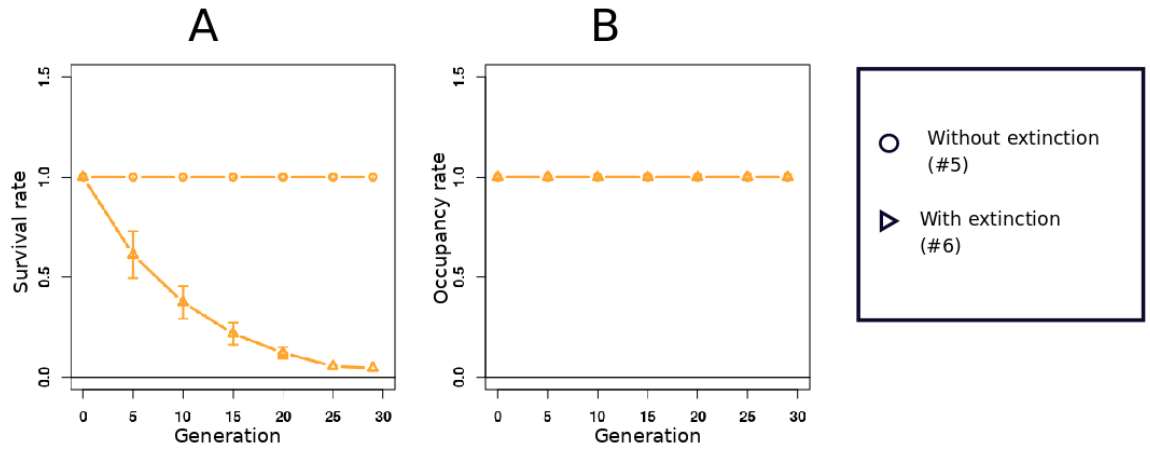


Figure 3: Survival rate (A) and occupancy rate (B) computed for the 50 populations of the metapopulations in cases with large size selfing populations without seed circulation, averaged over 10 replicates, according to the generation. Error bars correspond to the standard deviations.

#### 3.2.1.2 Results with seed circulation

Whatever the scenarios including seed circulation (scenarios **#9** to **#20**), the **survival rate** decreased during the first five generations and then remained almost stable over generations. As expected, this trend is intermediate with a lower survival rate compared to the situation without extinction (**scenarios #1, #3, #5, #7**) and a higher one compared to the situation with extinction but without seed circulation (scenarios **#2, #4, #6, #8**) (Fig. 4A, results are only shown for scenarios **#5, #6, #11, #15** and **#19** as the results remained the same for both large and small population sizes and selfing and outcrossing populations). This intermediate behaviour was roughly the same for all seed networks, reflecting the little impact of the network topology on the **survival rate**.

The **occupancy rate** remained globally unchanged whatever the scenario (Fig. 4B, results are only shown for scenarios **#5, #6, #11, #15** and **#19** as the results remained the same for both large and small populations size and selfing and outcrossing populations).

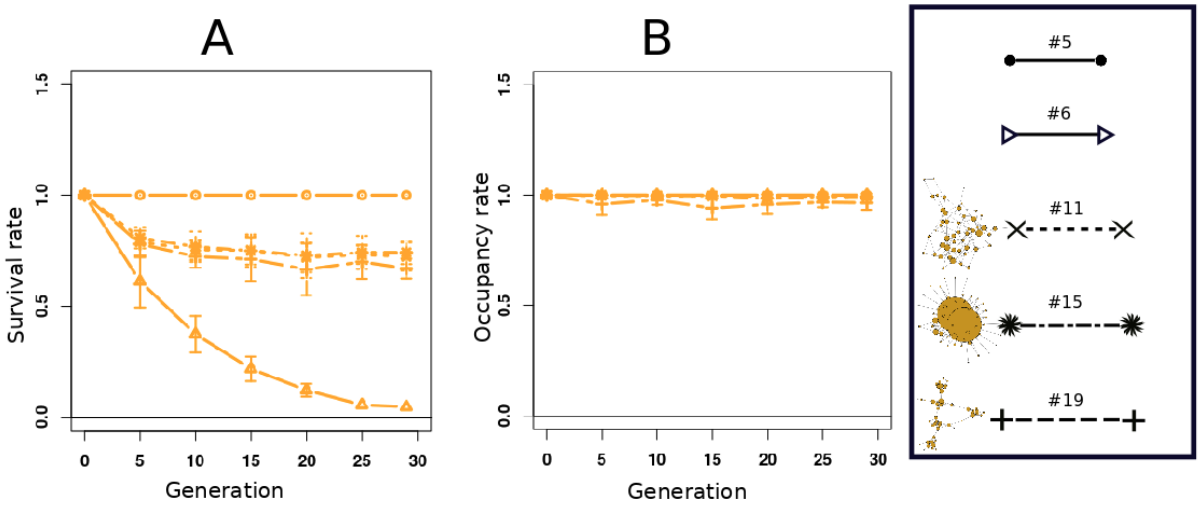


Figure 4: Survival (A) and occupancy rate (B) computed for the 50 populations of the metapopulation in cases with large size selfing populations with different types of seed circulation, averaged over 10 replicates, according to the generation. Error bars correspond to the standard deviations.

### 3.2.2 Genetic results

#### 3.2.2.1 Cases with large population sizes

**Genetic diversity** Whatever the scenario (**#5,#6,#7,#8,#11,#12,#15,#16,#19,#20**), the average within-population genetic diversity remained stable over generations ($H_s$ around 0.5). These results suggest that populations were large enough to avoid genetic drift and loss of genetic diversity (Fig. 5A ).

**Genetic differentiation** Whatever the scenario (**#5, #6, #7, #8, #11, #12, #15, #16, #19, #20**), the level of differentiation remained very low ($F_{ST}$ lower than 5%). The cases with large population size

have led to very little genetic drift and therefore to minimal genetic differenciation regardless of the process of extinction and colonization (Fig. 5B).

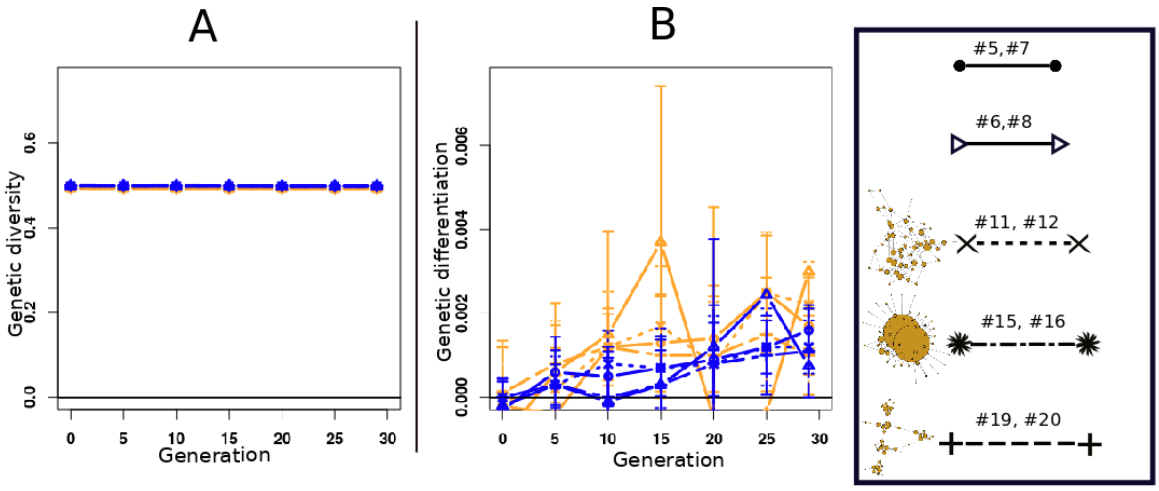


Figure 5: (A) genetic diversity and (B) genetic differentiation, averaged over 10 replicates, according to the generation for cases with large population sizes: blue color represents outcrossing populations, orange color represents selfing populations. Error bars correspond to the standard deviations.

#### 3.2.2.2 Cases with small population sizes

**Genetic diversity**

Whatever the scenario (**#1, #2, #3, #4, #9, #10, #13, #14, #17, #18**), genetic diversity decreased over generations ($Hs$ decreasing to 0.45 and 0.40 for the outcrossing and selfing crop populations, respectively) (Fig. 6A). These results differ from what was observed for cases with large population size. Note that the diversity decreased faster in the selfing crop populations than in the outcrossing ones because of the smaller genetic effective population size in selfing populations compared to outcrossing populations of the same demographic size. No impact of the type of network could be detected in either selfing or outcrossing crop populations. Decrease in genetic diversity in cases with small population size is expected as a consequence of genetic drift. The smaller the populations, the greater the sampling effect over generations.

**Genetic differentiation**

In all cases (**#1, #2, #3, #4, #9, #10, #13, #14, #17, #18**), genetic differentiation increased over generations to reach a rather high level between 9 and 24% (Fig. 6B). Differentiation was lower in the outcrossing populations (reaching a level of 9 to 13%) than in the selfing populations (17 to 24%) again because of the smaller genetic effective population size in selfing populations compared outcrossing populations of the same demographic sizes.

In both cases, there was a slight effect of the type of seed circulation network, larger for the selfing species. In particular, centralized networks (**#17,#18**) led to a lower differentiation among populations than the other networks, the cases without any seed circulation (**#1,#2,#3,#4**) showing the largest levels. In centralized networks, when farmers lose their seeds they often obtain new seed lots from the same central actor which may induce greater homogeneisation among farms.

This is consistent with what might be expected, since the less people exchange seeds, the greater the differentiation between populations (Fig. 6B). Moreover, the impact of the type of seed circulation network on genetic differentiation among populations is always more visible when populations are submitted to strong genetic drift pressures.

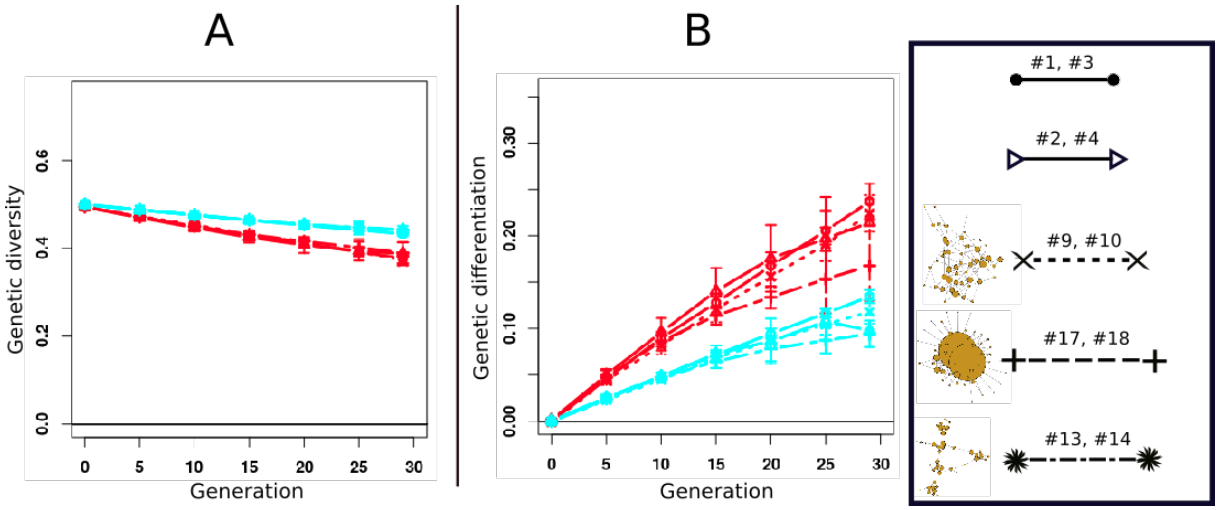


Figure 6: (A) genetic diversity and (B) genetic differentiation, averaged over 10 replicates, according to the generation for cases with small population sizes: blue color represents outcrossing populations, red color represents selfing populations. Error bars correspond to the standard deviations.

# 4 Learnings from and with community seed systems on their functionning

The following section shows the work conducted with the four CSSs (RSP tomato and wheat groups, RSR wheat and RAS tomato groups), with different levels of investigation ranging from describing their functionning to co-designing scenarios of crop diversity management, simulating the scenarios and assessing the impact of some key factors for the CSS (environmental conditions and farmers' practices, seed circulation network, extinction and colonization modalities,...). The approach followed an iterative process summarized in figure 7.

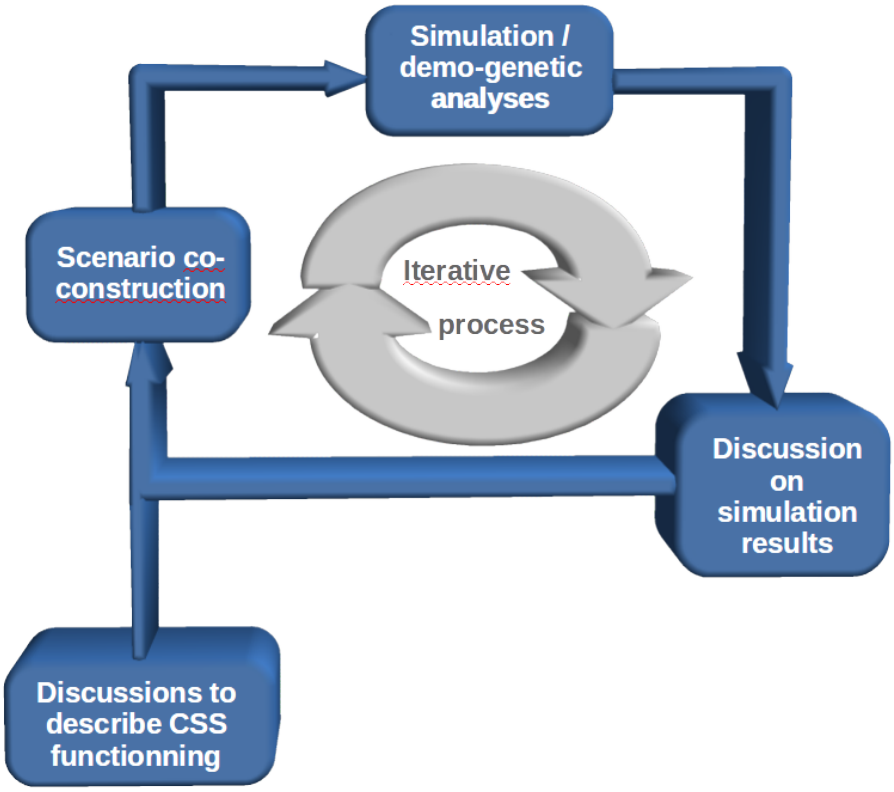


Figure 7: Description of the iterative process implemented in this DIVERSIFOOD activity.

## 4.1 Example of a community seed system managing bread wheat in Italy

**Initial condition**

Since 2011, the Rete Semi Rurali (RSR) is supporting the experimentation, selection and distribution of a bread wheat evolutionary population with a broad genetic basis (called "SOLIBAM population" after the EU project in which the population has been developed). Starting from two farmers in Sicily and Tuscany, the population spread to more than 50 farmers in different regions of Italy (fig. 8). Every years, some SOLIBAM populations are evaluated together with local and modern varieties and mixtures, in two farms where they are rated by farmers, technicians and consumers. After several years of evolution, some of the SOLIBAM populations, in particular those from Sicily and Tuscany, have become adapted to their growing conditions. They also appeared quite stable and received a high score from the farmers.

**Changes**

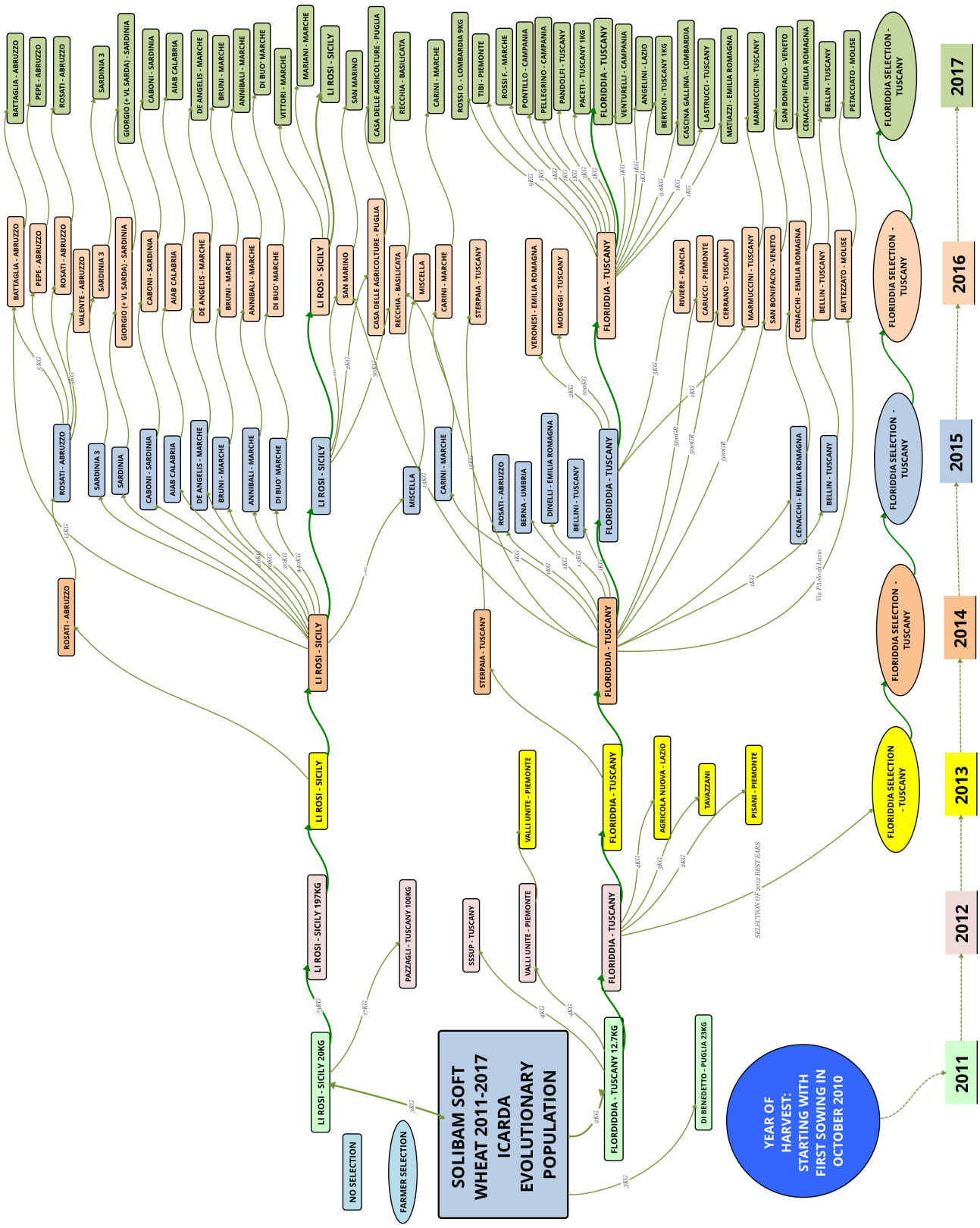


Figure 8: History of the SOLIBAM population from 2011 to 2017 (credits: Rete Semi Rurali)

Due to its great agronomic interest shown in the on farm evaluations, the SOLIBAM population has been proposed "heterogeneous material" for the experiment of marketing heterogeneous material. This has recently led to a larger scale distribution of the populations.

**Specific question addressed**

As the different populations derived from the SOLIBAM population have sometimes been evaluated as drastically different from each other, and have patterns of local adaptation, the issue has been raised on how the seed distribution to new farmers should be organized to optimize the adaptation and adaptability of the populations maintained in the CSS ?

**Progress**

Since 2017, the modelling team (INRA & CIRAD) participated in three meetings (during DIVERSIFOOD annual meetings) and in several phone meetings with RSR facilitators and in one meeting in Italy with farmers. Seed circulation among participants of the project have been recorded since the beginning of the project in 2011 (fig. 8). This information was used to estimate the extinction rate, the colonization rate and the seed network. Geolocalization of most of the farms was provided by RSR's facilitators and was used to collect additional data associated to the climate and the soil properties of each location.
Based on this information three different ways to represent the environments where the wheat populations have been grown were investigated: (i) all locations were considered to be submitted to the same environmental conditions (i.e. a single environment), (ii) the locations were assigned to two contrasted environmental conditions depending on the origin of their seed lot (i.e. two environments), (iii) the locations were assigned an environmental index value computed from the pedo-climatic data collected (i.e. environmental gradient) (fig. 9).

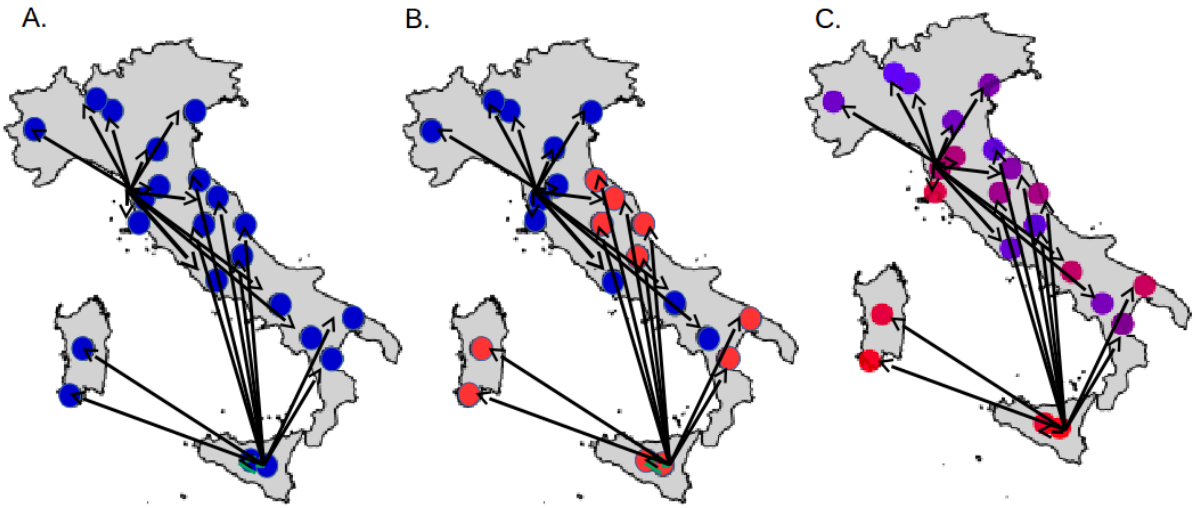


Figure 9: Three contrasted ways to represent the environments and their consequences in terms of selective pressure: A) same environment everywhere in blue; B) two contrasted environments: blue and red; C) a gradient of environments ranging from blue to red.

Simulation results show that genetic diversity ($H_s$) was maintained for the gradient of environments wheareas it decreased when populations evolved in only one or two different environments with seed circulating only among farms located in the same environment (fig. 10A.). In terms of differentiation ($F_{ST}$), a higher value was observed for the two different environments whereas one environment did not show any differentiationthe and the gradient of environments showed intermediate pattern (fig. 10B.).

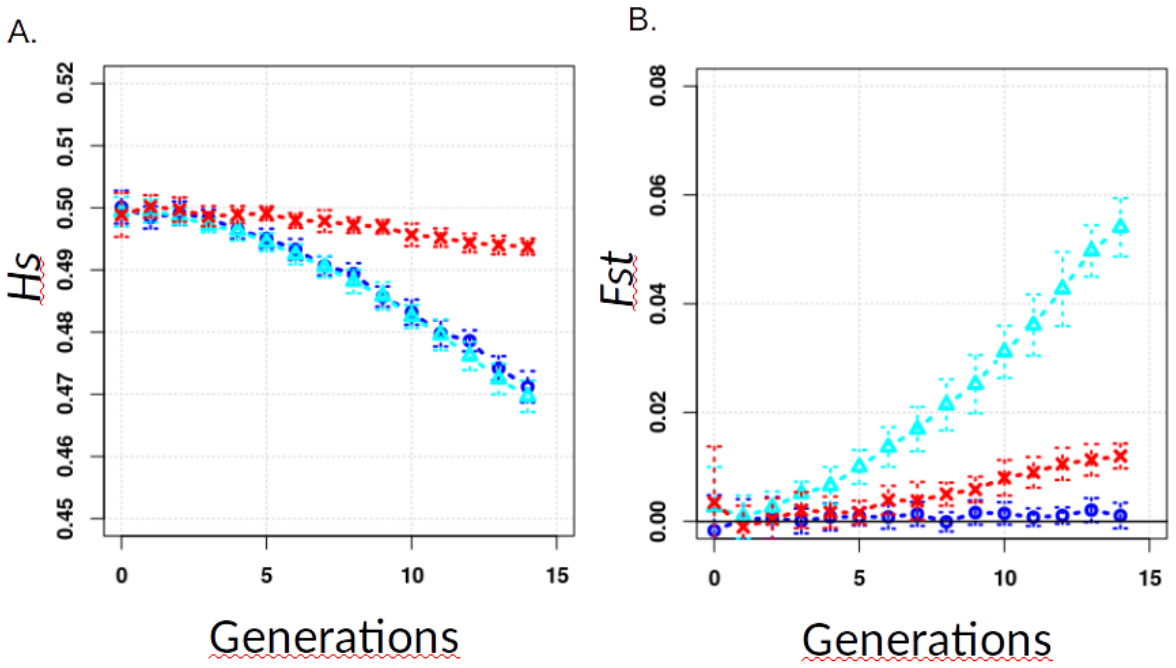


Figure 10: Simulation results for : A) genetic diversity ($H_s$), B) genetic differentiation ($F_{ST}$). Blue line corresponds to one environment; cyan line corresponds to two different environments; red line corresponds to the gradient of environments.

## 4.2 Example of a community seed system managing bread wheat in France

**Description of the CSS**

A historical Participatory Plant Breeding project has been conducted from 2006 to 2018 with farmers, farmer-bakers, farmer-millers of the Réseau Semences Paysannes (RSP) and research organization (INRA). The RSP wheat group is organised in local or regional collectives that are members of the RSP. The local organization within each collective follows particular rules while they also follow some rules designed within the PPB project among all members based on shared values. The main aims of the PPB project are to develop new populations stable over time and adapted to the local conditions of each farm and to increased farmers and facilitators autonomy with regards to seeds and crop diversity management. The reproduction of populations, their selection, mixtures, crosses within farms and the seed circulation among farms has been recorded over time and is stored in a dedicated database (SHiNeMaS). Moreover, agronomic and morphological characterization of a large number of populations has been carried out and is also stored in the database. The number of farmers involved as well as the number of collectives has increased over time starting with a single farmer in 2006, around 10 in 2008 and to around 100 in 2018 distributed in 7 collectives. This has led to issues in organising meeting involving all actors and in finding the time to share information from all collectives.

**Initial condition**

National scale of the project, interactions among researchers and farmers mainly through the RSP facilitators and little impact of the local / regional collectives.

**Changes**

Relocation of the seed circulation and of interactions among the participants at the local / regional scale. New practices have appeared among the participants of the program (creation and selection of mixtures).

**Specific question addressed**

Two main issues have been raised within the CSS: What is the impact of a more local circulation of seeds and of a differentiation among collectives in the management practices? To what extent does mixture of wheat populations impact crop diversity evolution in the wheat CSS?

**Progress**

Seed circulation among participants of the PPB project and some information on the environmental conditions and farming practices are available thanks to the information stored in the database (fig. 11).

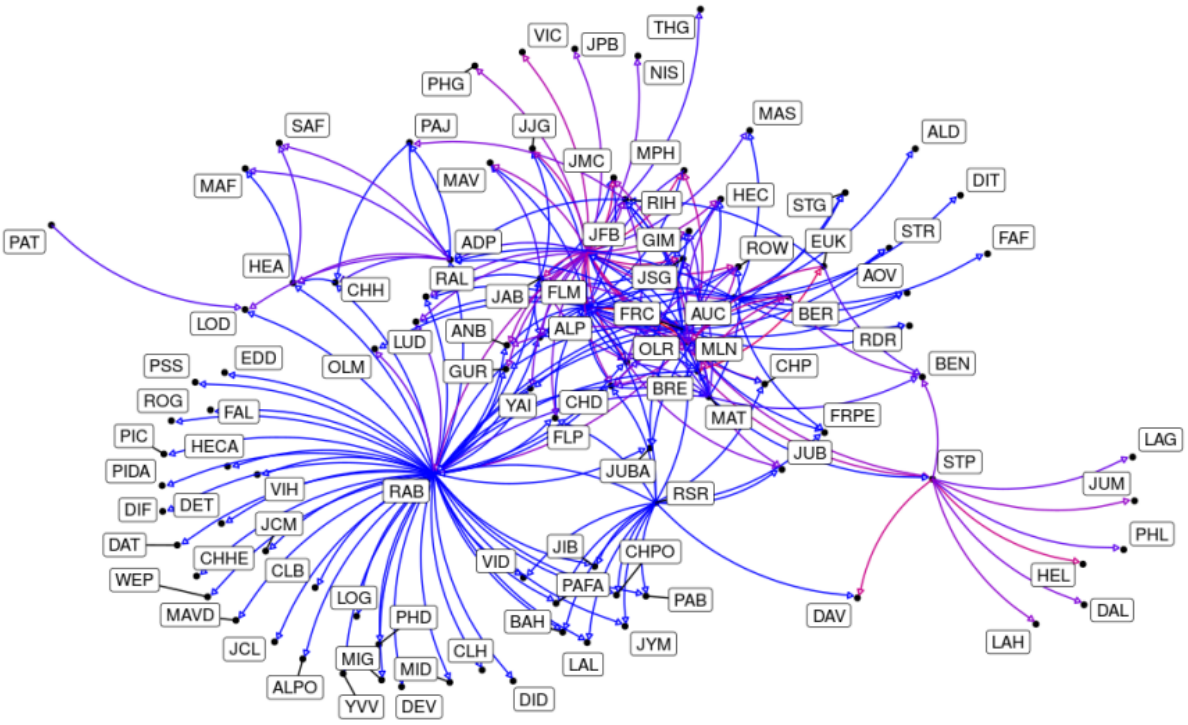


Figure 11: Seed network of the PPB program of bread wheat in France.

## 4.3 Example of a community seed system managing tomatoes in France

**Description of the CSS**

Graine del païs (GdP), member of the Réseau Semences Paysannes (RSP), is a small scale seed company created in 2005 localized in the Sourth of France specialized in multiplication and selection of landraces of vegetables and flowers. GdP works with around 30 seed producers mainly located in the South of France (fig 12). GdP sells seeds to several hundreds of gardeners every year. During the process of describing and co-designing scenarios of crop diversity management within DIVERSIFOOD, the focus has been put on the production of tomatoes seeds.

**Initial condition**

GdP manages all reference seed lots of all varieties. At each step of seed multiplication, GdP asks one of the seed producers to multiply part of this reference seed lot. The harvest is then sent back to GdP (fig 13 A: star strategy).

**Changes**

GdP is considering reorganizing the seed multiplication and conservation in order to better value the management work of each producer. Each of the 30 seed producers of tomato landraces will become responsible

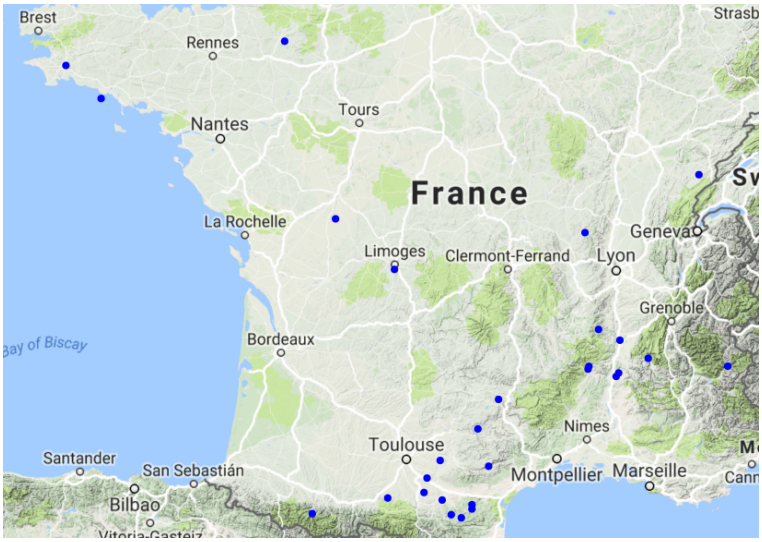


Figure 12: Map representing the locations of the seed producers involved in Graine del Païs seed multiplication in blue dots.

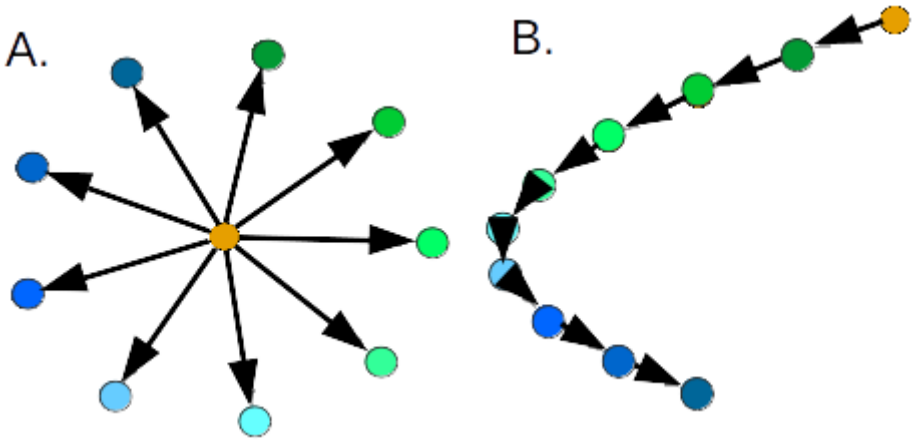


Figure 13: Description of Graine del Païs organization in seed multiplication: A) star strategy; B) stepping stone strategy.

of one or more tomato landraces, managing the reference seed lots of this landrace, asking others to multiply part of his reference seed lot (fig 13 B: stepping stone strategy).

**Specific question addressed**

Actors of the GdP collective are very interested in the issue of local adaptation. Therefore, they are wondering to what extent will this new collective organization impact adaptation and adaptability of the landraces maintained in the CSS ?

**Progress**

Since 2017, the modelling team (INRA & CIRAD) participated in two meetings with the GdP coordinator and in one meeting with the seed producers. The discussions helped the researchers to understand the collective organization and its on-going evolution. Circulation events, seed quantities and location of the producers and gardeners are recorded in a database since 2010. This information needs to be extracted and

will be used as input data for the model. Scenarios will be co-constructed and assessed during a PhD project that started in October 2018.

## 4.4 Example of a community seed system managing tomatoes in Andaluza, Spain

### Description of the CSS

The Red Andaluza de Semillas (RAS) is a regional organization member of the national Red de Semillas. RAS manages a Community Seed Bank (CSB) in Sevilla that mainly focuses on vegetables, in particular on tomatoes. The CSB is composed of facilitators and members: farmers, market gardeners or amateur gardeners mostly located in Andaluza (fig 14). The tomato seed circulation has been registered in a database since 2010.

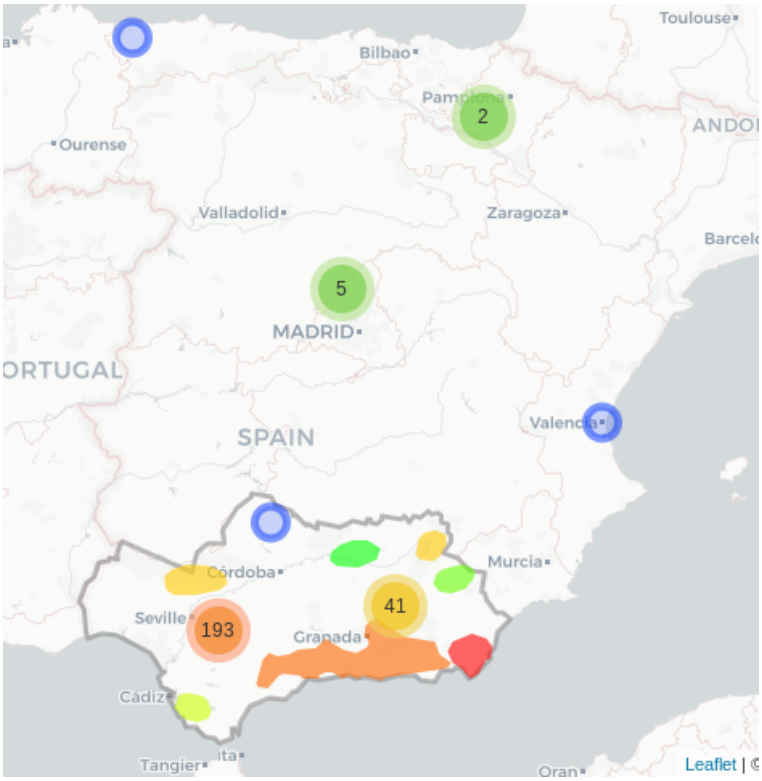


Figure 14: Map representing the location of the seed givers and receivers who participated in seed fairs organized by RAS since 2010.

### Initial condition

Only one CSB located in Sevilla manages seed ciculation at the Andaluza scale.

### Changes

Recently, new local CSBs have emerged. Now, the CSS is composed of different CSBs leading to : a) an increase in the number of actors involved in the CSS ; and b) new pathways for seed flows within and between CSBs.

### Specific question addressed

To what extent does this change in the CSS social organization have an impact on the crop genetic diversity maintained in the system ?

**Progress**

Since 2017, the modelling team (INRA & CIRAD) participated in three meetings (during DIVERSIFOOD annual meetings) and in one meeting with RAS coordinator to describe the functionning of the CSB. Access to the database was also provided. First information was extracted, analysed and presented (fig. 15). Now, the first scenarios need to be defined and assessed through simulation.

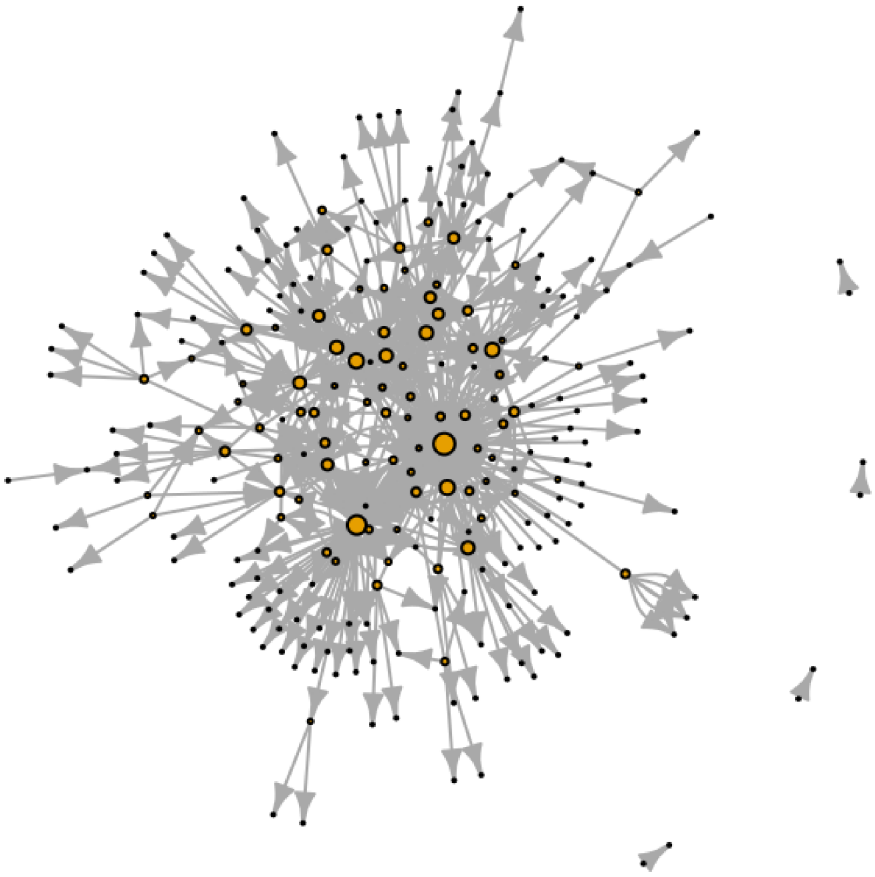

Figure 15: Seed network of Red Andaluza de Semillas since 2010.

# 5 Conclusion

## 5.1 Main messages

### 5.1.1 From theoretical approach

The theoretical case studies showed that seed circulation among actors (corresponding here to the introduction of a new seed lot on a farm after the loss of the population) allows to maintain the survival rate of the populations close to one and therefore helps to avoid the extinction of the metapopulation in the long term.

At the genetic level, in the cases with large population size (at least 10000 plants per population), the type of seed circulation network has had little impact on within-populatory genetic diversity or on the genetic differentiation among populations. This was due to the fact that these indicators remained quite stable due to the very limited impact of genetic drift. However, for cases with a smaller population size

(e.g. tomatoes or cabbage for which less than 100 plants per population are generally grown by farmers), genetic drift reduced significantly within-population genetic diversity and increased dramatically among populations genetic differentiation. In these cases, the type of seed circulation network did not change within-population genetic diversity but significantly reduced genetic differentiation. Centralized networks appeared to have the strongest homogenizing effect probably due to a limited number of actors distributing seeds to everyone.

### 5.1.2 Learning from the case studies

In the real case studies experience, interviews with farmers and farmers' organization facilitators were essential to describe the general trends and specificities from one CSS to another. Although the approach was conducted with four CSS located in three different countries and managing two different crops, the results consistently showed that all in CSS, very diverse and contrasted local conditions were met in terms of both environments and practices. By contrast, a similar general dynamics of seed circulation was observed among the CSS, with the organization changing from star network to community network when the number of actors increased. For the more advanced case study in Italy, first modelling results highlighted a trade-off between the two processes : heterogeneity of local conditions vs organizational changes.

## 5.2 Recommendations

### 5.2.1 Recommendations for research

This activity confirms how community seed systems are really complex and adaptive systems. An iterative process is necessary to translate this complexity into relevant modelling settings. Four case studies, corresponding to four different community seed systems, have been covered during DIVERSIFOOD. This work highlights : 1) the various ways of organizing collectively to manage crop diversity according to the species and the country ; 2) the different issues that these CSS are facing. Nowadays, there is no theoretical framework in population genetics able to account for so much complexity. Forward simulation is a first step to develop an integrated theoretical framework. Now, development of statistical methods are necessary to help compare different scenarios and real data collected in the fields (genetic and phenotypic data).

Recommendation #1 : Collaboration between modelleing scientists, farmers and farmers' organization facilitators are essential to properly describe social and evolutionary processes happening in such systems.

Recommendation #2: These findings encourage to increase the number of case studies to capture a broader diversity of initiatives at the European scale and even at a larger scale in order to provide a more realistic picture of what is going on in on-farm management of crop diversity.

Recommendation #3: Theoretical developments in statistics are essential to make the link between co-constructed scenarios and read data.

### 5.2.2 Policy recommendations

This participatory approach based on co-designing scenarios of community seed systems functionning showed that the collective dimension of CSS is key to seed circulation which allows for crop diversity development and maintenance. Therefore, this is important to support the collective organizations involved in the management, conservation and breeding of crop genetic diversity. These findings are in line with Diversifood outcomes which show that: 1) CSS contribute to reinforce the link between citizens and agriculture; and 2) the deployement of crop diversity in the fields is crucial for sustainable landscape, agriculture and food.

Recommendation #1: The second pilar of Common Agriculture Policy should include in the article promoting biodiversity the financial support for collective organizations. This support should be conditional on the type of diversity they manage, the number of people involved in the organization and their networking activities (number of meetings and events).

Recommendation #2: In the context of the compliance of the multi-lateral system (ITPGRFA), countries are invited to register the actors of the management of crop diversity (holders of plant genetic resources). Description of these actors should not focus only on the type and the amount of resources they hold and their accessibility but also if actors are collective organizations or individuals. In the first case, the actors should confirm they have networking activities.

Recommendation #3: Regions and municipalities should implement long-term financial support to local community seed systems.

# 6 References

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press.

Weir, B. S., and C. C. Cockerham. 1984. "Estimating F-Statistics for the Analysis of Population Structure." *Evolution* 38 (6): 1358–70. https://doi.org/10.2307/2408641.

# 7 Appendix

## 7.1 Survival rate: $S(t)$

(1)

$$S(t) = \frac{N_{pop}(t)}{N_{farm}}$$

where $N_{pop}(t)$ is the number of crop populations at a time $t$ and $N_{farm}$ is the total number of farms.

## 7.2 Occupancy rate: $O(t)$

Occupancy rate of population $i$ at time $t$:

(2)

$$O_i(t) = \frac{n_i(t)}{k_i}$$

where $n_i(t)$ is the number of plants in farm $i$ at a time $t$, and $k_i$ is the carrying capacity of farm $i$.

Average occupancy rate of the metapopulation at time $t$:

(3)

$$O(t) = \frac{1}{N_{pop}(t)} \sum_{i=1}^{N_{pop}(t)} O_i(t)$$

where $N_{pop}(t)$ is the number of crop populations in the metapopulation at time $t$.

## 7.3 Genetic diversity: $H_s(t)$

In each population, the expected heterozygosity of population $i$ at time $t$, $H_s^i(t)$, is estimated at the locus level and averaged over all the loci as follows (Nei 1987):

(4)

$$H_s^i(t) = \frac{n_i(t)}{n_i(t) - 1} \frac{1}{L} \sum_{l=1}^{L} (1 - \sum_{k=1}^{K} p_{i,k,l}(t)^2)$$

where $L$ is the number of loci, $K$ is the number of allele per locus, and $p_{i,k,l}(t)$ is the allelic frequency of allele $k$ at locus $l$ in population $i$ at time $t$, and $n_i(t)$ is the number of plants in farm $i$ at time $t$.

(5)

$$H_s(t) = \frac{1}{N_{pop}(t)} \sum_{i=1}^{N_{pop}(t)} H_s^i(t)$$

where $N_{pop}(t)$ is the number of loci at time $t$.

## 7.4 Genetic differentiation: $F_{ST}(t)$

The genetic differentiation is computed at the metapopulation level following the variance decomposition:

(6)
$$F_{ST}(t) = \frac{\sigma_a^2(t)}{\sigma_a^2(t) + \sigma_b^2(t) + \sigma_w^2(t)}$$

where $\sigma_a^2(t)$ is the variance among populations, $\sigma_b^2(t)$ is the variance among individuals within population, $\sigma_w^2(t)$ is the variance among genes within individual.

This variance decomposition is used to compute the $F_{ST}$ estimator $\theta$, developped by Weir and Cockerham (1984) and used in this study to estimate genetic differentiation.